Criptography

Tutorial #5

Manuel Barbosa (`mbb@fc.up.pt`) Rogério Reis (`rogerio.reis@fc.up.pt`)

MSI/MCC/MERSI — 2022/2023

1. In the file `data.py` the list `packets` contains 1000 messages (portuguese ASCII) enciphered by a stream cipher. We know that some of those messages were enciphered using the same key stream. Can you identify them?

2. Alice and Bob agree to communicate privately via email using a scheme based on **RC4**, but they want to avoid using a new secret key for each transmission. Alice and Bob privately agree on a 128-bit key k. To encrypt a message m, consisting of a string of bits, the following procedure is used.

   i - Choose a random 80-bit value v.

   ii - Generate the ciphertext $c = RC4(v \cdot k) \oplus m$, where $\cdot$ denotes the concatenation.

   iii - Send the bit string $(v \cdot c)$.

   (a) Suppose Alice uses this procedure to send a message m to Bob. Describe how Bob can recover the message m from $(v \cdot c)$ using k.

   (b) If an adversary observes several values $(v_1 \cdot c_1)$, $(v_2 \cdot c_2), \ldots$ transmitted between Alice and Bob, how can he determine when the same key stream has been used to encrypt two messages?

   (c) Approximately how many messages can Alice expect to send before the same key stream will be used twice?

   (d) Write a **Python** program that, given a size n, computes the smallest number of uniformly random generated numbers, $(r_i)_i$ (such that $0 \leq r_i < n$), for which it is more likely to have a repetition (in the generated numbers) than not.

   Hint: Start by writing a closed formula for the number of functions

   $$f : A \longrightarrow B,$$

   where $|A| = k$ and $|B| = n$. Then, write a similar closed formula, but for the injective functions $f : A \to B$.

   What is the probability of having a "collision" for a given k and n?

   If the numbers are very large you will probably need to use the Stirling approximation to a factorial:

   $$\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}e^{\frac{1}{12n}}.$$

An approximation is, thus,

$$n! \simeq \sqrt{(2\pi n)} \left(\frac{n}{e}\right)^n .$$

Although this obviates some cumbersome computations, the result still envolves intermediary computations too large to be carried out directly.
The probability of a collision, given k and n, can be approximated by

$$1 - e^{\frac{k-k^2}{2n}} .$$

But the search for the minimal value of k that makes the probability of having a "collision" above $\frac{1}{2}$, cannot be achieved in "brute force" mode. A better approach must be pursued...

(e) How many messages should Alice use the key k, before generating another?